FUNDS FOR LEARNING
BRINGING TECHNOLOGY TO THE CLASSROOM

RECEIVED

OCT - 2 2001

FEDERAL COMMUNICATIONS COMMISSION
OFFICE OF THE SECRETARY

October 2, 2001

DOCKET FILE COPY ORIGINAL

Magalie Salas
Secretary
Federal Communications Commission
445 12th Street SW
Room TW-A325
Washington, DC 20554

RE: In the Matter of Request for Review of the Decision of the Universal Service Administrator by Prince George's County Public Schools Under FCC Docket Nos. 97-21 and 96-45 (SLD Form 471 No. 199306)

Dear Ms. Salas:

Enclosed please find the original and four copies of the Request for Review of the Prince George's County Public Schools in the above-referenced matter.

Sincerely,

Orin R. Heend

Funds For Learning, LLC
2111 Wilson Blvd. #700
Arlington, VA 22201
703-351-5070

No. of Copies rec'd 0+4
List ABCDE

Before the
FEDERAL COMMUNICATIONS COMMISSION
Washington, DC

RECEIVED

OCT - 2 2001

FEDERAL COMMUNICATIONS COMMISSION
OFFICE OF THE SECRETARY

In the Matter of:	)	
	)	
Request for Review of the Decision of the Universal Service Administrator by	))	
	)	
Prince George's County Public Schools	)	SLD Form 471 No. 199306
	))	
Federal-State Joint Board on Universal Service	)	CC Docket No. 96-45
	)	
Changes to the Board of Directors of the National Exchange Carrier Association, Inc.	))	CC Docket No. 97-21

To: The Common Carrier Bureau

REQUEST FOR REVIEW

Prince George's County Public Schools ("PGCPS" or "School District"), by its representatives, hereby seeks review of the determination of the Schools and Library Division of the Universal Service Administrative Company ("SLD"), dated September 4, 2001, denying in part the request of PGCPS (Billed Entity No. 126359) for universal service support in Funding Request Numbers 472448, 472445, 472403, 472390, 472354, 472345, 472342, 472177, 472172, 472155, and 472153, and denying in full its request in No. 472417. (SLD Form 471 No. 199306).

I. Facts

On September 4, 2001, the SLD issued a Decision on Appeal in connection with the above-referenced matter. (Attachment A). The SLD decided, in pertinent part, not to approve universal service support for the School District to purchase Dell caching servers. In support of its decision, the SLD reasoned as follows:

> Caching servers are storage devices for Internet content, and in accordance with program rules Internet content is an ineligible product/service. The SLD has determined that Caching servers are ineligible for discounts under this program.

The SLD's conclusion that caching servers are content "storage devices" and thus ineligible for support under the Commission's definition of "Internal Connections" is wrong. Moreover, it is not even consistent with the Commission's current position, which is that caching servers "seem to provide levels of efficiency in the delivery of information, but are not necessary to transporting such information." *Request for Review by the Department of Education of the State of Tennessee of the Decision of the Universal Service Administrator, Request for Review by Integrated Systems and Internet Solutions, Inc., of the Decision of the Universal Service Administrator, Request for Review by Education Networks of America of the Decision of the Universal Service Administrator,* CC Docket Nos. 96-45 and 97-21, Order, 14 FCC Rcd 13734 (1999) ("*DOE Tennessee*") at para. 41.

For the reasons set forth below, we request the Commission to reconsider and reverse its determination in *DOE Tennessee* that caching servers are ineligible for universal service support, to grant the Request for Review on that basis, and to instruct the SLD to approve and fund, to the extent that funding is available, all of the requests in issue.

II. Issue

Whether Caching Servers are necessary to transport information within one or more instructional buildings of a single school campus and thus eligible for universal service support.

III. Discussion

A. Reconsideration of the caching server eligibility issue is warranted because the original determination was not based on a complete and unbiased record.

Based on the record before it in *DOE Tennessee*, the Commission concluded that a type of file server known as a caching server failed to satisfy the definition of "Internal Connections" and thus was ineligible for universal service support. Under Section 54.506 of the Commission's Rules, a service is an eligible internal connection if it "is necessary to transport information within one or more instructional buildings of a single school campus." As noted above, the Commission found that caching servers "seem to provide levels of efficiency in the delivery of information, but are not necessary to transporting such information." We disagree. A caching server is an intelligent device that is actively involved in transporting information to and from the local area network ("LAN"). It is a "traffic cop," essentially a switch that controls the amount and type of new, web-based information that will be allowed at any given time to enter the LAN.

For the reasons set forth below, we submit that in today's bandwidth-taxed school and library LAN environments, caching servers are not only critically necessary to transport information, they are the last line of defense against network congestion that at any time can bring the flow of information from the Internet to a complete grinding halt.

The central caching server issue in *DOE Tennessee* was whether the servers' cost could be eligible for support as part of the underlying facilities necessary to provide end-to-end Internet access. Unfortunately, the other caching server issue, *eligibility,* arose in an ancillary context that did not lend itself to a full, impartial briefing of the issue. That was because it was not in any party's best interest to convince the Commission that caching servers should be eligible as "internal connections." Indeed, the parties knew that their respective and collective interests would be far better served if the Commission concluded just the opposite. *See DOE Tennessee* at para. 39 ("...when the rules of priority are in effect, there is an incentive to characterize certain facilities used in the provision of internal connections that may also be provided by the Internet access service provider as Internet access service").

At the time *DOE Tennessee* was decided, not enough E-rate support remained to fund the caching servers in issue as Priority Two Internal Connections. Both parties knew, therefore, that if the Commission decided that caching servers satisfied the "internal connections" test, their ultimate objective would be more difficult to reach. That objective was to convince the Commission to fund the caching servers instead as a component of Priority One Internet Access. They knew that the Commission would be less likely to grant support for a caching server as a component of Priority One Internet Access, if the exact same "component" was eligible for support as a Priority Two Internal Connection. Thus, both the applicant and its service provider stood to benefit greatly if the Commission concluded that caching servers were *not* internal connections. Neither party, therefore, had any incentive to flesh out the facts. Accordingly, we urge the Commission to reconsider the issue, taking into account the discussion and information set forth below, as well as in the attached documentation. (*See* Attachment B).

B. <u>Caching servers are eligible Internal Connections because they are necessary to transport information within one or more instructional buildings of a single school campus.</u>

1. Caching Servers receive requests for web-based information and, in response to each request, decide whether to (a) retrieve and deliver new information from the originating web server; (b) deliver information stored in cache; or (c) deliver a combination of both.

The combination of increased Internet usage and media rich web sites has created an almost insatiable demand for bandwidth on school-based LANs. Because network caching provides an economical solution to this growing concern, school districts are turning increasingly to caching servers to help keep their LANs up and running. The object of Internet caching is to transport information to the network's edge, to store it there temporarily, and then to deliver it quickly to end-users upon request. Caching increases network *speed* (decreased response times), *throughput* (the network can accommodate more students and teachers simultaneously), and *availability* (web site accessibility more likely even if the originating web server is down).

Internet caching is a multi-step process. First, the caching server monitors the network for web-based information requests or URLs. When it receives one, the server checks it against the web sites that it already has in cache (memory). If the site is there, its next job is to implement a validation process to determine whether the information is current. If it is, there is a "hit," and the caching server immediately will direct a copy of the web site to the computer that requested it. If, on the other hand, the server determines that the web site needs refreshing and/or contains dynamic content (e.g. stock quotes, sports scores, news) that needs to be downloaded to keep the site fresh, the caching server will retrieve the current information from the Internet, copy and store all or part of it in cache, and deliver another copy to the end-user.

Each hit on the caching server eliminates a long round trip across numerous Internet hops to the originating web server and back. Hits dramatically increase network response time and available bandwidth (*speed*). This in turn minimizes the likelihood of network paralyzing congestion, enables more users on the LAN to request the same information simultaneously and successfully, and allows substantially more information to be downloaded over the same period of time (*throughput*). Occasionally, it even affords end-users access to web sites that are down or for some other reason temporarily inaccessible *(availability)*, making it even more evident that caching servers must be Internal Connections.

2. Caching servers are "local theaters" of the Internet.

To understand just how critically important caching servers are to transporting information from the Internet to classrooms, consider this analogy.[1] What would happen if the motion picture industry had no distribution system, no local theaters? What if everyone who wanted to see a Hollywood movie had to travel to a single Hollywood theater to see it? There would be massive congestion. Traffic would be backed up for miles. Eventually, the line of people at the theater would grow so long that even the most enthusiastic moviegoer would have no choice but to give up and return home disappointed. But this of course never happens because the motion picture industry devised a simple solution – use a local distribution network to move creative content closer to the audience. By distributing copies of films to local theaters everywhere, the industry eliminates bottlenecks that would occur if the movie were to play at a

1 *See* Intel, *Caching for Improved Content Delivery, Bringing the Web Closer with Caching Technologies*, included as a part of Attachment B.

single location. This of course makes it easier and faster for the public to get to those films and enables producers to reach a substantially larger audience.

Caching servers are the Internet's "local theaters," as they are the local distribution points in the Internet's global information distribution system. Upon request, caching servers contact web servers throughout the world, copy the static, and in some cases dynamic, content they find there, transport it back to the local area network, deliver it to the original requestor and, thereafter, for a pre-programmed period of time, make it available to other local users -- making it easier, faster, and cheaper for everyone on the LAN to access the identical information -- and without the cost of popcorn.

3. Caching server functionality results from integrating closely together several different independent functions to create a whole that is substantially greater than the sum of its parts.

Characterizing a caching server as a content storage device completely misses the point as to what this hardware is designed to do. Simply because a sophisticated piece of equipment performs a storage function does not automatically turn it into a storage device. For example, does an aircraft carrier store aircraft or transport them? Obviously, this engineering marvel does both, yet one would never dream of calling it a storage ship. Aircraft land on its deck, where they remain - in storage - while in transport to a site that is much closer to the aircrafts' ultimate destination. This enables the aircraft to reach their targets much more quickly. Caching servers, of course, transport, store and deliver information in much the same fashion. Thus, even though the two very different examples of hardware both have storage functions, in neither case is that its main or only function. In short, the storage label fails utterly to reflect the sophisticated, integrated nature and purpose of each.[2]

2 Numerous other eligible Internal Connections have storage functions that are integrated into the hardware's overall design. Basic network file servers, for example, have an extremely large storage capacity, as do web servers. Indeed, caching and web servers perform almost identical functions, except that caching servers are far more "intelligent" and more directly involved in the transport of information.

4. A caching server provides levels of efficiency in the delivery of information...... but so does a switch.

In *DOE Tennessee*, the Commission decided that caching servers were not Internal Connections because they "seem to provide levels of efficiency in the delivery of information, but are not necessary to transporting such information." We submit that the Commission failed under the circumstances to assign sufficient weight to the "level of efficiency" that caching servers actually provide, and that caching servers, like switches, provide a level of efficiency in delivering information that make them vitally "necessary" to the transport of information in all but the smallest networking environments.

No one would dispute that a switch satisfies the test for Internal Connections, yet peer to peer local area networks can and do operate quite well *without* one. Does that mean that a switch is not "necessary" to transport information? Of course not. As local area networks get larger and more sophisticated, it would be unthinkable not to incorporate switches into the network design. In those networks, switches provide a level of efficiency in the delivery of information that would make it virtually impossible as a practical matter to deliver that information without them. This applies equally to caching servers. To attempt to transport information over a LAN of any size in today's high bandwidth environments without a caching server would be similarly unthinkable.

As the Commission has seen fit to categorize as Internal Connections switches and other hardware that are not in every instance literally "necessary" to transport information, we can only conclude that the Commission has decided to adopt a broader, more practical, common sense definition of the term.[3] Accordingly, we urge the Commission to apply that definition to caching servers, as this category of network electronics certainly is as necessary and, in some cases, even more necessary to transporting information within schools than many of the items

---

3 Multiplexor equipment is yet another example of an eligible internal connection that is not "literally" necessary to transport information. Like a caching server, however, it makes delivering that information much less expensive. Another example is UPS (uninterruptible power supply) equipment. That eligible hardware is much like a caching server in that it plays a critically important role in making sure that information continues to move across a LAN without interruption.

currently listed on the FCC/SLD Eligibility List.

5. Caching servers satisfy the Internal Connections test for file servers that the Commission set forth recently in *Cleveland Municipal School District.*

In contrast to caching servers, there clearly are servers, such as CD and database servers, which are ineligible storage devices. Unlike caching servers, however, which hold data only *temporarily* and act as "intelligent" conduits for information on its way to and from the LAN, CD and database servers are simply *permanent* repositories of information.

In *Cleveland Municipal School District*,[4] the Commission held that Student Data Warehouse servers failed to satisfy the test to determine whether a file server is eligible for funding as internal connections because "the servers act as the source of content, not as conduits for content which originates elsewhere." [5] No one would dispute that the content that is stored temporarily on caching servers originates elsewhere. Further, it is undisputable that caching servers are conduits for this information, as that is precisely their job -- to retrieve and distribute information from Internet web servers and, thereafter, to distribute it to anyone on the LAN who requests it. Therefore, under the eligibility test that the Commission announced in *Cleveland Municipal School District*, caching servers must be eligible for E-rate support as Internal Connections.

4 See *Request for Review of Cleveland Municipal School District*, CC Docket Nos. 96-45, 97-21, Order, File No. SLD-190883 (Com. Car. Bur. Rel. Aug. 16, 2001)
5 *Cleveland Municipal School District* at para. 8.

IV. RELIEF SOUGHT

PGCPS requests that the Commission reverse the SLD's funding decision with respect to caching servers, approve funding for all of the Dell caching servers for which support has been requested, and instruct the SLD to fund fully, where possible in view of available universal service support, all of the FRNs in issue.

Respectfully submitted,

PRINCE GEORGE'S COUNTY PUBLIC SCHOOLS

By: ______________________

Orin R. Heend
Funds For Learning, LLC
2111 Wilson Blvd. Suite 700
Arlington, VA 22201
703-351-5070

cc: Roland Moore
Chief Information Officer
Prince George's County Public Schools
14201 School Lane
Upper Marlboro, MD 20772

USAC **Universal Service Administrative Company**
Schools & Libraries Division

Administrator's Decision on Appeal - Funding Year 2000-2001

September 4, 2001

John D. Harrington
Funds For Learning
Re: Prince George's County Schools
229 North Broadway
Edmond, OK 73034

Re:	Billed Entity Number:	126359
	471 Application Number:	199306
	Funding Request Number(s):	471985, 471989, 472029, 472054, 472448, 472445, 472403, 472390, 472354, 472345, 472342, 472177, 472172, 472155, 472153, and 472417
	Your Correspondence Dated:	August 25th, 2000 (2 Letters)

Attachment A

After thorough review and investigation of all relevant facts, the Schools and Libraries Division ("SLD") of the Universal Service Administrative Company ("USAC") has made its decision in regard to your appeal of SLD's Year Three Funding Commitment Decision for the Application Number indicated above. This letter explains the basis of SLD's decision. The date of this letter begins the 30-day time period for appealing this decision to the Federal Communications Commission ("FCC"). If your letter of appeal included more than one Application Number, please note that for each application for which an appeal is submitted, a separate letter is sent.

Funding Request Number:	471985, 471989, 472029, 472054
Decision on Appeal:	**Approved in full**
Explanation:	

- Your appeal has brought forward persuasive information that discounts should be provided for these requests.

Funding Request Number:	472448, 472445, 472403, 472390, 472354, 472345, 472342, 472177, 472172, 472155, and 472153
Decision on Appeal:	**Partially Approved**
Explanation:	

- Caching servers are storage devices for Internet content, and in accordance with program rules Internet content is an ineligible product/service. The SLD has determined that Caching servers are ineligible for discounts under this program. Your documentation indicates that the caching servers make up 21.3% of these funding requests. It has been determined that the digital tape and HP OpenView had no associated cost, and should not have been removed from your request. Consequently these requests will be approved less the cost of the caching servers.

Funding Request Number: 472417
Decision on Appeal: **Partially Approved but Denied for Funding**
Explanation:

- Your appeal claims that caching servers should be considered eligible because their function is similar to that of web servers. You also feel that it is unfair that the eligible services list on the SLD web-site did not identify caching servers as ineligible. You claim that the digital tape and HP Open View included in the request are eligible since they are bundled with the servers with no associated cost.

- Caching servers are storage devices for Internet content, and in accordance with program rules Internet content is an ineligible product/service. The SLD has determined that Caching servers are ineligible for discounts under this program. Your documentation indicates that the caching servers make up 21.3% of these funding requests. It has been determined that the digital tape and HP Open View had no associated cost, and should not have been removed from your request.

- However, for Funding Year Three, there are not sufficient funds to provide internal connections discounts to applicants at this discount rate. On your Form 471 you indicated that the entity receiving this service has a discount eligibility of 40%. Consequently, SLD denies your appeal because there is insufficient funding for Funding Year Three to provide discounts for internal connections requests to applicants that are below the 82% shared discount level.

- FCC rules require that where demand for funding exceeds available support, first priority be given to requests for telecommunications services and Internet access. *See* 47 C.F.R. §54.507(g)(1)(i). FCC rules further require that requests for internal connections be given second priority, and be funded only if funds remain after support has been provided for telecommunications and Internet access through all discount levels in a funding year. *See* 47 C.F.R. § 54.507(g)(1)(ii). Where demand for discounts for internal connections exceeds available support, FCC rules require funding be awarded first to applicants eligible for a ninety percent discount level, and then at each descending single discount percentage until funds are depleted. *See* 47 C.F.R. § 54.507(g)(1)(iii).

Since the Administrator's Decision on Appeal approves additional funding for your application, SLD will issue a new Funding Commitment Decision Letter to you and to each service provider that will provide the services approved for discounts in this letter. SLD will issue the Funding Commitment Decision Letter to you as soon as possible. The Funding Commitment Decision Letter will inform you of the precise dollar value of your approved funding request. As you await the Funding Commitment Decision Letter, you may share this Administrator's Decision on Appeal with the relevant service provider(s).

If you believe there is a basis for further examination of your application, you may file an appeal with the Federal Communications Commission, Office of the Secretary, 445 12th Street, SW, Room TW-A325, Washington, DC 20554. Please reference CC Docket Nos. 96-45 and 97-21 on the first page of your appeal. Before preparing and submitting your appeal, please be sure to review the FCC rules concerning the filing of an appeal of an Administrator's Decision, which are posted on the website at <www.universalservice.org>. **You must file your appeal with the FCC no later than 30 days from the date on this letter for your appeal to be filed in a timely fashion.**

We thank you for your continued support, patience, and cooperation during the appeal process.

Schools and Libraries Division
Universal Service Administrative Company

CC: Michael Lieb

Attachment B

Caching for Improved Content Delivery

Bringing the Web Closer with Caching Technologies

intel

Table of Contents

Executive Summary

The enormous success of the Web as a source of information and a platform for e-Commerce has not come without challenges. One ever-increasing problem is the highly variable and often frustrating length of time it takes to access a web site and download pages, prompting cynics to claim that "www" stands for "World Wide Wait." While this presents serious challenges to both marketer and consumer alike, it also presents opportunities.

Carriers and service providers are making huge investments to increase Internet bandwidth. However, by itself, additional bandwidth cannot address network latency or accelerate slow origin servers. Enter caching. This technique addresses the challenges of the Web by moving content closer to users who need it. Caching has immediate benefits not only for the end user, but also for Internet service providers and content providers. And, in a future where every business is an e-Business, it can give any site a major competitive advantage.

This paper looks at the case for Web caching, providing an overview of caching technology and the implementation requirements. It also describes various deployment options, identifies ideal cache locations and considers the strengths of caching appliances. Additional resources on caching are provided at the end of the paper.

Why Implement Web Caching?

Meeting the Internet Challenge

The explosive growth of the Web has severely stressed Internet capacity and performance. Carriers and Internet service providers have responded with massive investments to expand capacity, from the Internet backbone to the "last mile" into businesses and homes.

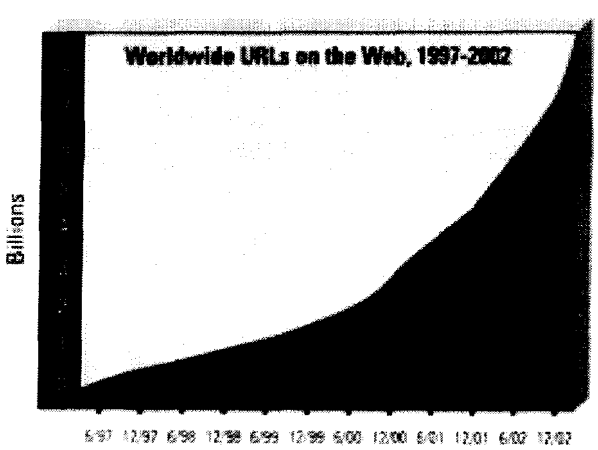
However, both the number of Web users (Figure 1) and the amount of Web content accessed are predicted to accelerate dramatically in the years ahead. The Web is becoming a major center for business transactions of all types, with an increasing proportion of traffic taking the form of e-Commerce. The projected value of this rapidly expanding Internet economy is enormous. The growing importance of the Web as a stimulant for economic growth means that it must become a more reliable and predictable place to do business.

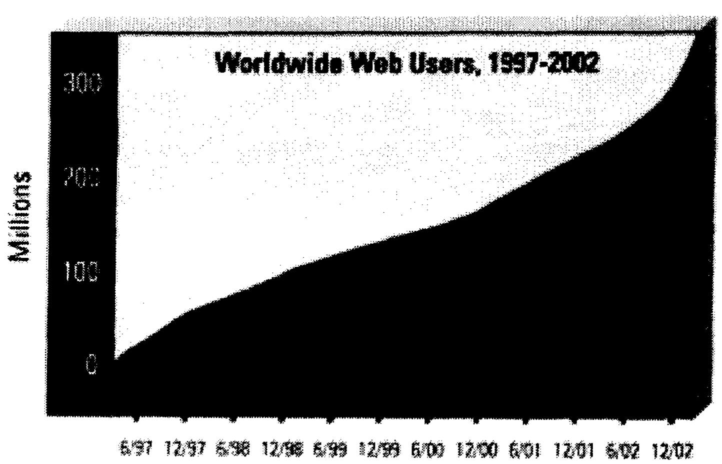
Currently, the Web is fundamentally inefficient. Every user seeking to view specific content must obtain it directly from the server that is the point of origin for that content. This is the equivalent of having everyone fly to Hollywood to see the latest movie. There is no distribution mechanism designed into the Web that is analogous to the system of movie theaters that offer first-run films in every viewer's hometown.

Since it is not possible to have dedicated, point-to-point bandwidth allocated to users, congestion is inevitable. Problems contributing to user frustration include:

- Slow connection speeds
- Unpredictable performance
- Limitations in available bandwidth
- Overwhelmed Web sites

Figures 1 and 2: Growth, both in the number of Web users and the amount of content those users are accessing, is predicted to accelerate after the turn of the century. This will place a tremendous load on the Internet and potentially impact users' quality of service experience. Source: International Data Corporation, 1998





The capacity of the Internet is constantly being built out to handle the growing load. For the foreseeable future, this build-out will continue to lag behind demand. In any case, simply increasing bandwidth by building up the network with bigger pipes cannot address all of the Quality of Service (QoS) issues involved. For purposes of this discussion, QoS means a high quality user experience, measured in low latencies for downloads and fast download times. Adding bandwidth may improve speed, but not latency or delay. In addition, adding bandwidth at one point may only move a bottleneck to another location.

Caching makes more bandwidth available by using existing pipes more efficiently, not only improving QoS for the user, but also giving service providers substantial savings and additional room to grow (for details, see "Who Benefits?" below).

What is Caching?

Caching is a technology that is already familiar in other applications. Many hardware devices cache frequently used instructions and data in order to speed processing tasks. For example, data that is frequently used by a computer's central processing unit (CPU) is stored in very fast memory, sometimes right on the CPU chip, thereby reducing the need for the CPU to read data from a slower disk drive. In addition, Web browsers are designed to cache a limited amount of content on a user's PC. That is why selecting "Back" or "Previous page" on a browser toolbar typically results in near-instantaneous retrieval.

With true Web caching, the same concept is applied more widely, using a server or specialized appliance. Web content is placed close to users in the form of a network cache, reducing the number of routing/switching hops that are required to retrieve content from a remote site. In other words, viewers aren't required to travel to Hollywood to see a movie, rather movies are sent to local theaters where people can access them—or better yet, the viewers themselves determine which movies are made available locally.

There are two kinds of Web caching models. In the "edge-services" model, businesses subscribe to a third-party service vendor to have their content cached. This has serious disadvantages for some of the parties:

- The ISP doesn't own or control the infrastructure.
- The most frequently used sites are not necessarily the ones cached, which is a disappointment to users.

In the open model, supported by Intel caching appliances, service providers install their own caching equipment and are able to offer caching as an added-value service to their customers. Advantages include:

- The ISP invests in its own destiny, not that of a third party.
- Additional revenue can be realized directly by the service provider.
- The system automatically caches the sites that users access most.

Who Benefits from Caching?

End users, user enterprises, service providers and content providers all stand to benefit substantially from caching implementation.

The ultimate beneficiaries are end users, the people who drive the Internet economy. Caching provides distinct benefits for end users in the form of an enhanced Internet experience and better perceived value for their monthly service fees.

Caching also has benefits for enterprises. By providing a local cache for Web content, companies can monitor how much Internet bandwidth is required to satisfy employee needs. They can also initiate access policies to limit employee use of the Web to corporate activities.

For Internet service providers, caching has several important advantages:

- It reduces Internet bandwidth usage by eliminating redundant requests for popular documents.
- Leased line expenses are reduced or postponed. Benchmarks have shown that if a cache successfully serves a modest percentage of user requests, the amount of outbound bandwidth required can be reduced by up to 30–40%. That can mean a significant cost savings, or the ability to add more users with the current network. To access a bandwidth calculator for individual computations, please see **http://www/intel.com/network/products/cache/nscache.htm.**
- Caching also provides better QoS, leading directly to higher customer satisfaction and reduced customer turnover. Less spending is needed for acquiring new customers.
- A caching solution provides the basis for new value-added Web hosting services that boost ISP profitability.

Content providers benefit from higher site availability and a better user experience with fewer, shorter delays. This creates increased customer satisfaction, giving cached sites a competitive advantage over those that are not cached.

Studies indicate that a delay of only five to eight seconds is enough to frustrate the average user into retrying or leaving a site. Caching helps prevent this. And, from an overall commercial viewpoint, users can visit more sites, do more shopping and purchase more products if content can be delivered and downloaded faster.

Overview of Caching Technology

With and Without a Solution in Place

Without a caching solution in place, requests for content from a browser and the content delivered from the origin server must repeatedly take the same long-distance trip—from the requesting computer to the computer that has the content, and back (Figure 3). The following steps are typical:

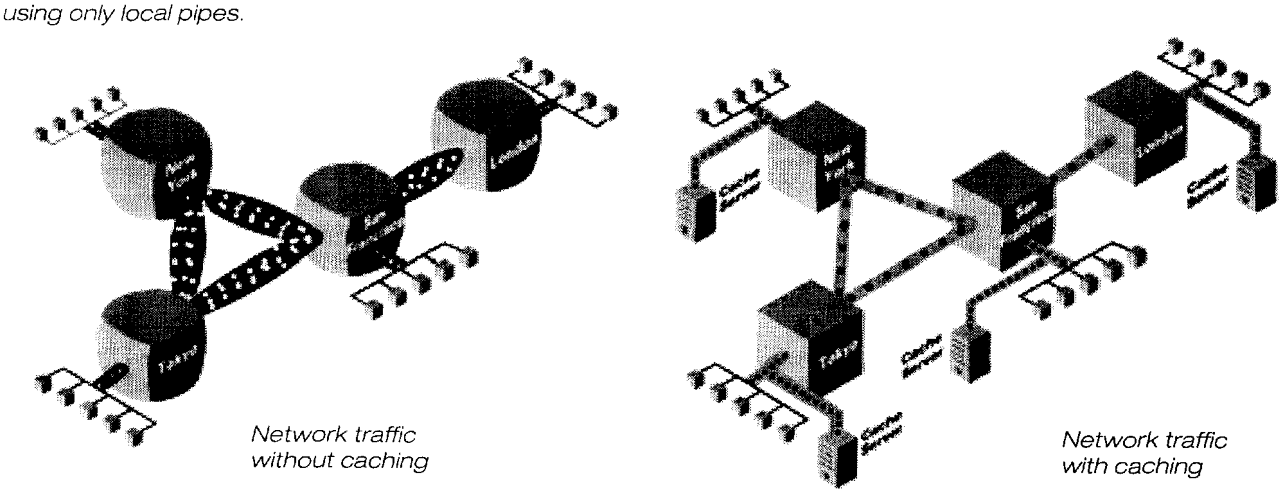
- The Web browser sends a request for a Uniform Resource Locator (URL) that refers to a specific Web document on a particular server on the Internet.
- This request is routed through the normal TCP/IP network transport.
- Content requested from the server (also known as an HTTP server) may be a static HTML page with links to one or more additional files, including graphics. The content may also be a dynamically created page that is generated from a search engine, a database query or a Web application.
- The HTTP server returns the requested content to the Web browser one file at a time. Even a dynamically created page often has static components that are combined with the dynamic content to create the final document.
- If there is no cache server in place, the next user who requests the same document—even if that user is in the next cubicle—must send a request across the Internet to the originating Web server and receive the content by return trip.

When caching is used, the process is far more efficient because frequently accessed content does not have to repeatedly make the long trip from the origin server (Figure 3).

- The requested document may be stored on a cache server inside the user's corporate LAN, at the user's ISP, or at some other Network Access Point (NAP) or Point of Presence (POP) located closer to the user than the majority of Web servers.
- If the requested document is stored on the cache server, then the server will check to make sure the content is current (fresh). To ensure that a user does not receive a stale object, freshness parameters are pre-set by content providers and others, and servers are normally configured with default algorithms.
- If the content is fresh according to these parameters, then the transaction is considered a cache "hit," and the request can be immediately fulfilled from the cache server.
- If the content needs to be refreshed, the cache server will retrieve updated files from the Internet and send them to the user, also keeping fresh copies for itself.
- The more frequently a cache can serve user requests, the higher the hit rate and the better the performance enjoyed by users.

Similar processes are involved for FTP file transfers, with an FTP server handling each request for a file submitted by the FTP client

Figure 3: The amount of bandwidth required for trips across the backbone is significantly greater in a non-cached network. With caching configured, a large portion of the requests can be fulfilled using only local pipes.



application. Delays and bottlenecks can be an even bigger problem with FTP because the size of a typical FTP file is larger than a typical HTML file. Streaming audio and video are additional examples of Internet applications that can benefit from caching content. Internet latency problems can cause jittery video and delayed or distorted audio. Better use of bandwidth can be a solution for these problems.

Reducing Bandwidth Usage

Along with giving users an improved experience, caching reduces the upstream bandwidth an ISP has to provide to fulfill user content requirements. A cache only passes user requests on to the Internet if it isn't able to service them. The greater the number of user requests that can be fulfilled from a cache, the less bandwidth is used to reach distant origin servers. This traffic reduction means significant savings for a service provider, since an estimated one-third of an ISP's operating costs are recurring telecommunications charges.

It's true that freshness updates must be performed, so there would still be traffic from the ISP out to the Internet even if all requested content were to be found in the cache server. But by using caching, bandwidth utilization can be greatly reduced. Caching is even beneficial when retrieving dynamic documents, because these pages do have some static elements that can be served from a cache.

Depending on the distribution of traffic and the scalability of the cache, up to 40% (Source: Patricia Seybold Group, 1999) of user HTTP requests can be taken off the network and fulfilled from the cache server. This makes networks far more efficient,

Key Requirements

The most important requirement of a caching solution is the ability to provide optimized performance. There are two sides to cache performance:

Operational capacity: This is addressed by the architecture and the implementation of the cache server. Along with raw cache capacity, architectural issues include how the server makes use of multiple threads of execution, and how well it performs load balancing in a multiple cache server deployment.

Responsiveness to user requests: This is determined by the various techniques the cache server uses to maximize hit rate, including the structure of hierarchies (see "Cache Hierarchies," below) and content optimizations. Cache hit rate is a function of many factors, including the cache size and the load on the cache.

A cache server can be tuned to improve capacity and responsiveness in many ways. Potential areas of optimization include:

- Processing queues for the various objects that make up a document
- Determining whether or not a requested object is cached
- Delivering the requested object to the browser if it is not cached
- Total throughput based on how fast incoming requests are handled

Performance depends on how well these possibilities were understood and used by those who built the cache server and engineered the software.

Scalability for the Web is another key requirement that a cache must address. The effectiveness of caching improves as the traffic served by the cache increases—the bigger the challenge, the more valuable the solution. To support very large caches, cache server clustering or load balancing is necessary, and the caching solution must support these capabilities.

Cache server support is also required for a variety of protocols. Network caching can be applied to content delivered over HTTP, NNTP, FTP and others. All are characterized by having some proportion of static content.

Manageability is mandatory for any caching solution. Cache management includes the ability to easily install and maintain cache servers, and to access the wealth of usage and traffic data the servers can provide. It is often necessary to manage arrays of cache servers, which may be distributed over great distances, from a single point of control.

Browser-based management interfaces are increasingly common as the standard way to manage distributed systems. The interface should provide functionality for configuring cache servers, administering security, setting filters, loading the cache, controlling the cache system and gathering information from logs.

A caching solution should be designed to provide high reliability and availability. Although the duplicative nature of caching has a measure of fault tolerance built in, the solution must feature high quality software and a highly reliable platform if it is to be considered an integral part of the network infrastructure.

Configuration approaches such as failover and clustering also contribute to reliability and availability.

Finally, hardware and software must be well integrated to achieve the efficiencies that are at the heart of caching performance.

Deployment Models

There are several approaches, or models, for implementing a cache architecture. Which model is chosen depends on where the cache is implemented, the primary purpose of the cache and the nature of the traffic.

Forward Proxy

Forward proxy cache is defined by its reactive nature. With a forward proxy cache configuration, user requests go through the cache on the way to the destined Web server. If the cache contains the requested document, it serves it directly. If it does not have the desired content, the server acts as a proxy, fetching the content from the Web server on the user's behalf.

Reverse Proxy

A cache can also be configured as a fast Web server to accelerate slower, traditional Web servers. Documents stored in cache are served at high speed, while documents not in cache—usually dynamic content and other short-term objects—are requested when necessary from the origin Web servers. This model is frequently used to optimize the performance of a Web server site. The caching system sits in front of one or more Web servers, intercepting requests and acting as a proxy.

Cache servers can be deployed throughout a network to create a distributed network of sites for hosted content, a model that is sometimes referred to as site replication. In addition to performance benefits for the user and content provider, reverse proxy caching also has benefits for the ISP. Those benefits include the ability to enable load balancing, to offer peak-demand availability insurance and to provide dynamic mirroring for high availability.

Transparent Caching

Forward proxy caches can be further configured as either transparent or non-transparent. A transparent cache sits in the network flow and functions invisibly to a browser. The benefits of caching are automatically delivered to clients without anyone having to reconfigure browsers. For ISPs and enterprise backbone operations, a transparent configuration is often preferred because it minimizes the total administrative and support burden. Individual users and small businesses without IT staff also appreciate the absence of configuration requirements.

The most popular implementation is to use a Layer 4 capable switch to interface cache servers to the Internet (Figure 4). These switches can inspect network traffic and make decisions above the IP level. For example, the switch can direct HTTP (or other) traffic to the cache and send the rest of the traffic directly to the Internet. The switch can also send requests to specific nodes in a cache server cluster, a capability that can be used for load balancing purposes. Using a pair of switches with multiple cache servers allows for redundancy and failover protection.

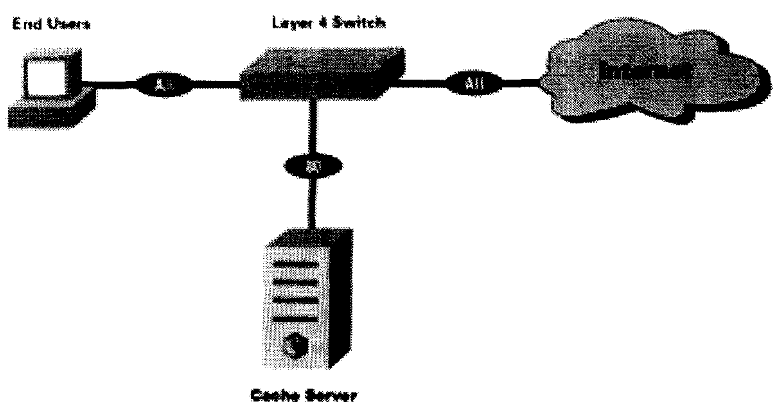
Cache Locations

To identify ideal cache deployment points, there are three types of location characteristics to keep in mind:

Choke point. Traffic convergence points or choke points are locations where a large majority of network traffic passes and would therefore be visible to a cache server. This allows the cache to handle more requests and store more content than if located somewhere that is easily bypassed.

High traffic load. Any area characterized by high traffic conditions allows higher cache utilization. The more cache hits, the greater the benefits.

Figure 4: Layer 4 switches can route requests for cacheable data (HTTP, NNTP, etc.) to the cache server while sending other requests to the Internet.



Economic potential. Points where users will benefit from high cache hit rates while also reducing upstream bandwidth requirements will provide both QoS benefits and positive economics for the access provider.

These characteristics are typically found at major Internet switching locations, dial-in aggregation points, or corporate gateways (Figure 5). Uses include standard POP and dial-up access, NAPs and exchanges. Web hosting, "last mile" acceleration, satellite-based cache feeding and more. Caching is even employed as an economical means of updating information for online news services.

Cache Hierarchies

In the event that a requested document is not stored in cache (a cache "miss"), the cache server usually must forward the request to a distant origin server. However, if the cache server were able to check with another nearby cache instead, the process could be much faster. This is the idea behind cache hierarchies.

It is possible to create relatively small regional caches—for example, a server or cluster handling a department or limited geographical area—and link them to larger parent caches that define larger groups or areas. If a regional cache does not have a requested document, it can forward the request to the parent cache. This will still provide faster service than contacting the origin server. Multiple-level hierarchies can be configured, giving cache servers a sequence of larger and larger caches to query if the first attempt misses.

By combining capabilities such as site replication and a linked hierarchical caching structure, a highly efficient distributed network can be created for Web hosting over a wide geographical area.

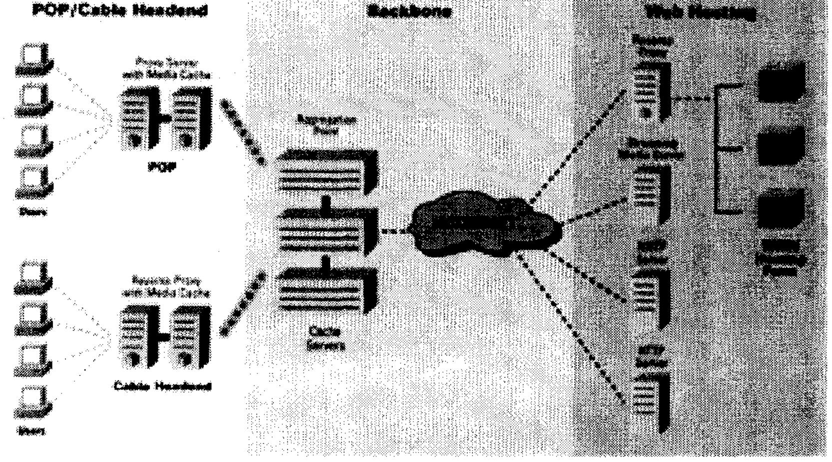


Figure 5: Cache servers may be placed at an ISP POP to serve requests locally, at an aggregation point on the edge of the Internet to reduce bandwidth requirements, or in front of a Web farm to reduce load on content servers.

Advantages of Using a Cache Appliance

While this paper is intended to provide information on Web caching in a generic context whenever possible, Intel® products are used in the section below in order to provide a meaningful level of detail. Unless otherwise indicated, the appliance functionality and attributes described below are applicable to the caching appliances offered by Intel.

Cost-effectiveness

By definition, an appliance (sometimes referred to as a "thin server") is a device that provides a limited number of dedicated functions, and is therefore able to deliver those functions more cost-effectively than a multi-purpose device. This doesn't mean that appliances are not robust solutions. In fact, by specializing in one particular area, they often provide a richer feature set, superior stability and broader flexibility in terms of deployment and configuration.

As example, the Intel® NetStructure™ Cache Appliance's integrated hardware and software design has been specifically engineered to provide robust, carrier-class caching. Capabilities include:

- Speed (the ability to handle thousands of simultaneous user connections)
- Scalability (nodes can easily be added as needed to a cache cluster)
- Fault tolerance (contributing to network stability)
- Secure single-point administration (many nodes can be configured at once)

Ease of Installation and Use

As a fully integrated "solution in a box" comprising all of the necessary hardware and software, an appliance is very easy to install and configure. The Intel NetStructure Cache Appliance has automated wizards and intuitive software configuration that make setup easy. This is a significant part of the cost savings provided by appliances, because it takes

Web Caching Solutions from Intel

As a network vendor, Intel supports Web caching with:

- **The Intel® NetStructure™ Cache Appliances:** Designed to provide a cost-effective method for improving bandwidth management and peak-load performance. These appliances feature industry-leading Inktomi Traffic Server Engine* software integrated with Intel technology on an Intel platform.
- **Intel® Express 550T Switch:** Designed for medium-sized businesses and service providers, the 550T is able to perform traffic management through Layer 4 redirection to a variety of Web caching servers, including the Intel NetStructure Cache Appliance. An economical alternative to high end Layer 4 switches.

minimal time to incorporate the device into the network and doesn't require the expertise and expense of a systems administrator.

Further savings are provided by the relatively compact size of most appliances. The Intel Net-Structure Cache Appliance, for instance, comes in a low profile, rack mountable design. This provides an easy way to increase network capacity in the same limited space that an infrastructure owner or operator already has available.

Flexibility

Since it is designed for a single, specialized purpose, an appliance typically offers a high degree of deployment flexibility. The Intel NetStructure Cache Appliance is no exception. It can be used in a variety of deployment models, alone or with other enterprise software, including other caching products. Here are some of the ways it can be implemented:

- Forward proxy
- Reverse proxy
- Transparent caching
- Part of an HTTP cache hierarchy
- ICP sibling: The Intel NetStructure Cache Appliance can send ICP queries to neighboring caches as part of an ICP cache hierarchy
- NNTP news cache: The Intel appliance caches frequently accessed news articles and can also receive news feeds for designated news groups

In addition, the Intel NetStructure Cache Appliance offers broad support for content and interoperability protocols:

- HTTP versions 0.9 through 1.1
- FTP
- NNTP
- ICP (to help implement cache hierarchies)
- SSL encryption
- WCCP
- WPAD

Performance

As previously noted, performance depends on capacity, including how well the server makes use of multiple threads of execution, and the ability to respond quickly to user requests. The Intel NetStructure Cache Appliance is designed for high-performance operation across a broad range of load conditions.

It aggressively implements multi-threading—breaking down large transactions into small, efficient tasks. A threaded event scheduler allows the Intel NetStructure Cache Appliance to handle thousands of simultaneous connections and maximize CPU usage. The appliance is able to respond to multiple requests simultaneously and efficiently even under peak loads.

The appliance's Inktomi Object Store is a custom-designed Web object database that has been fully optimized for caching. It uses raw disk I/O to achieve optimized storage and retrieval of content objects, resulting in much higher speeds than conventional file systems. In order to provide fast access for the most frequently requested objects, a RAM cache is maintained so that hot objects can be read from high-speed memory instead of from disk. In addition, all objects are indexed according to their URL and associated headers. This means the Intel NetStructure Cache Appliance can store, retrieve and serve not only Web pages, but also parts of Web pages, providing optimized bandwidth savings.

Centralized Administration

The Intel NetStructure Cache Appliance helps minimize the cost of administering, maintaining and operating a large cache system. It offers several centralized management alternatives to suit the needs of a wide range of environments:

Browser-based interface: The manager User Interface (UI) offers password-protected, single point administration for the entire Intel cache cluster.

Command line interface: A command line interface lets the administrator configure the system's network addresses, and control, configure and monitor the cache.

SNMP management: The Intel NetStructure Cache Appliance supports two management information bases (MIBs) for management through SNMP facilities. MIB-2 is a well-known standard MIB. A proprietary Intel Cache MIB provides more specific node and cluster information.

Performance reporting: Performance statistics are available at a glance from the manager UI or the command line interface. Some of the characteristics that can be managed include: log file formats; site or content blacklist filtering; anonymization; never-cache, pin-in-cache, revalidated after; store multiple versions of cached objects for user-defined or browser-defined differences in content; domain and host-name expansion, and content routing.

Scalability and Reliability

In recognition of the mission-critical nature of caching, the Intel NetStructure Cache Appliance is designed to provide a highly reliable and available cache service. And, since it is designed to implement caches at the highest levels of network traffic, including Network Access Points and on the backbone, it is easy to scale.

The Intel NetStructure Cache Appliance achieves a high degree of scalability through three mechanisms. They include:

- Symmetric Multiprocessing (SMP)
- Clustering
- Cache hierarchies

Multiple threaded processors provide the in-box performance to accommodate growth, and clustering provides scalability across several machines by spreading the workload. The Intel NetStructure Cache Appliance executes its own cache hierarchy configuration, which is used in conjunction with ICP to communicate with other caches.

Clustering technology is supported by the Intel NetStructure Cache Appliance, combining the resources of several machines to increase capacity. As new nodes are added to the cluster they build on existing nodes to provide additional disk and processing resources. Clustering also offers failover protection—node failures can be automatically detected, and traffic is then redistributed to active nodes.

Conclusion

In today's complex Web environment, it is important to consider end-to-end performance and response time as being the product of many factors, over which few Web sites, service providers or users have control. SPs need to provide an optimal user experience, measured in low latencies for downloads and fast download times.

Various caching approaches are available, and they can be implemented in a variety of ways depending on the specific caching requirements. When correctly placed and configured, caches can significantly improve the user experience and QoS, while saving service providers significant costs of providing upstream bandwidth. Another plus is the added revenue that caching can bring to SPs by giving them opportunities to offer service level guarantees and peak insurance.

The Intel NetStructure 1500 Caching Appliance, featuring Inktomi Traffic Server Engine* caching software, is a carrier-class product capable of delivering fresh content to a large number of users from a large number of Web servers. It is ideal for enterprises that need to better manage the use of network resources, provide superior information distribution to employees, and reduce the administrative burden through transparent proxy and caching capabilities. Even more importantly, it gives service providers a superior approach to managing growth in back-end connectivity—growth that otherwise could expand at an almost infinite rate.

For More Information

For more information about the Intel® NetStructure™ Caching Appliance, please visit **www.intel.com/network/products/cache/nscache.htm**

For complete details on the Inktomi Traffic Server Engine* web caching software, visit **www.inktomi.com**

To find out more about the Intel® Express 550T Switch with Layer 4 redirection for Web caching, see **www.intel.com/network/products/express_switches.htm**

POWERAPP.CACHE™ LEARNING CENTER

What is Internet Web Caching?

The concept of caching is not unique to the computer industry. Caches are collections of copied data. The idea of caching is used widely in computer system design. Microprocessors commonly use cache memories on the processor chip to access recently used data. Although this type of memory is relatively limited and comparatively expensive, the increase in speed provided by the cache is valued above these drawbacks.

Similarly, a web cache also stores recently accessed information. This cache is a dedicated computer system within the Internet that monitors web object requests, retrieves them, and then stores those objects. Subsequent users requesting the same objects (web sites) are served by the local cache instead of the web site's origin server. These cached objects relieve the need to go through multiple hops on the Internet route from user request to web site origin server and back. Otherwise, each of these hops along the route can cause delays in service.

The speed of any connection is always limited to the slowest link in the path. By keeping retrieved web objects closer to the users requesting them, web caches significantly enhance the speed at which those objects can be accessed and viewed. This local storehouse of objects (i.e., text pages, images, and other Internet content) is called the web cache.

Caching can also improve the overall efficiency and performance of corporate or ISP web hosts by front-ending the routine or repetitive requests for the relatively static information that make up 40 to 50 percent of web traffic, allowing web servers to concentrate on more dynamic content.

The Dell PowerApp™ .cache can be deployed in either Forward Proxy or Reverse Proxy mode. Forward Proxy is used to accelerate the Internet access response times of a LAN browsing out to the Internet. Reverse Proxy is used to front-end E-commerce and/or web Servers, effectively off-loading incoming requests for static content, and thus increasing the number of concurrent users/connections the web server is able to maintain, while at the same time, improving the browsing experience for those users pointing their browsers to the web server.

POWERAPP.CACHE™ LEARNING CENTER

Top Twelve Web Caching FAQs

Q **1. What is web caching?**

A A web cache is a dedicated system that monitors web requests and stores the content it retrieves from various web servers. On subsequent requests for the same object, the cache delivers the object from its local storage rather than passing the request on to the origin web server.

By handling object requests rather than passing them upstream to the origin server, caches reduce network traffic, relieve web server burden and improve the browsing experience for users. Caches can be located anywhere on a network, and each cache will store a different set of objects based on the needs of the users it serves.

The Dell PowerApp.cache can be deployed in either Forward Proxy or Reverse Proxy mode. Forward Proxy is used to accelerate the Internet access response times of a LAN browsing out to the Internet. Reverse Proxy is used to front-end E-commerce and/or web Servers, effectively off-loading incoming requests for static content, and thus increasing the number of concurrent users/connections the web server is able to maintain, while at the same time, improving the browsing experience for those users pointing their browsers to the web server.

Q **2. Why is caching important?**

A Web caching is a promising approach to the problem of rising Internet and Intranet traffic volume for two main reasons: quality of service and overall traffic reduction.

- **Quality of Service** - A cache located closer to the browser delivers frequently requested content through fewer routers, thus reducing the potential for packet loss delays and speeding overall service.
- **Traffic Reduction** - As a rule of thumb, caches have a "hit" rate of 35%, which means that 35% of content requested through them can be successfully cached, and they therefore reduce upstream traffic on the network by that same percentage.

Q **3. What types of caches are available?**

A Web caching products come in two forms: appliances and software.

- **Appliances**, like the Dell PowerApp.cache, integrate caching software with a hardware platform. These devices are designed for easy setup and require minimal administration.
- **Software** - Caching software products run on standard operating system platforms and server hardware requiring the customer to purchase each component separately and install the system like any other application.

Q **4. How is web caching different from browser caching?**

A Internet browser applications allow an individual user to cache web content as files on the user's local hard disk. A user can configure how much disk space should be devoted to caching web content. This method serves one user well, but does not benefit other users on the same network accessing the same web sites and therefore overall benefits of reduced WAN and server access traffic are negated. In contrast, shared network caching stores the content on a network-level device. When User A visits a web site, the content is cached on the local network. When Users B or C point to the same web page, the content can be provided from the local cache, improving response time and freeing up bandwidth that would have been used to access the web server and download the content.

Q **5. How does caching work with real-time content like stock quotes?**

A Many of the objects that make up a given web page are static and can therefore be cached. For example, banners and buttons are typically static objects. Other objects on a page can be dynamic and should not ever be cached for example, stock prices. With caches, only a few dynamic objects need to be retrieved from the origin server while the static objects are fulfilled locally. This method provides an effective way to increase performance while maintaining real-time data.

Q **6. How do I know that the content in the cache is up-to-date?**

A Any caching system must ensure that a user sees the same content from a cache as from the Internet or an intranet. Content freshness is controlled if the cache is HyperText Transfer Protocol (HTTP) 1.1 compliant. Web server providers using HTTP 1.1 have direct control over how long content is to be cached. The content author can set caching attributes for each object on a web page, as follows:

- Non-cacheable
- OK to cache (This is the default setting)
- Explicit time-out

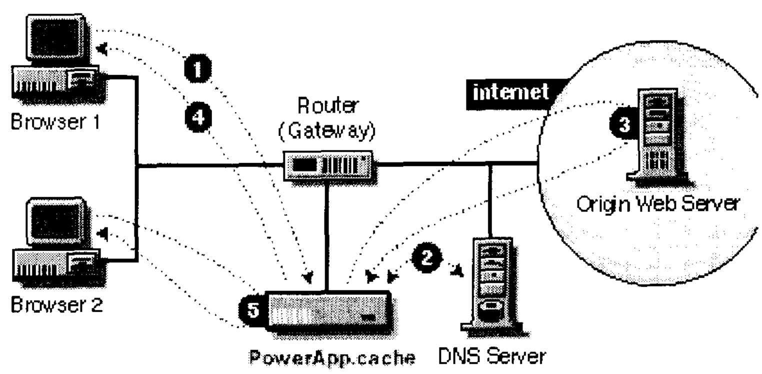
In addition, users can explicitly refresh content at any time by using the browser's reload button. All other data remains cached until it times out in the caching system.

Q **7. How are caches configured?**

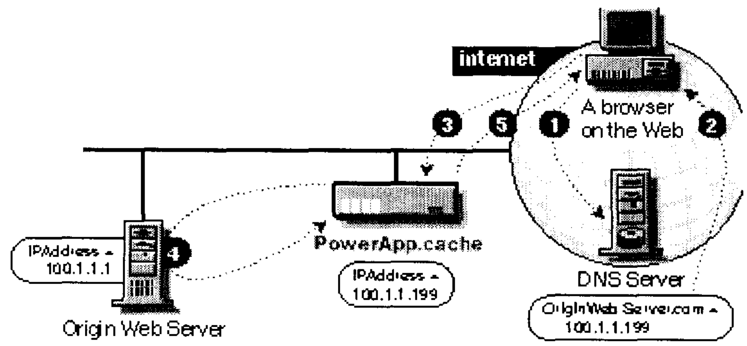
A

A cache can be configured as a proxy for browser users, as a transparent cache browser users, or "reverse proxy" caches.

- **Forward Proxy Cache** - A proxy cache operates by explicitly cooperating browser by directing HTTP requests to the cache rather than the target w The cache then either satisfies the request itself or passes on request to as a proxy for the browser (hence the name). Proxy caches are particular on enterprise Intranets, where they serve as a firewall that protects intran against attacks from the Internet. The most obvious disadvantage of the configuration is that each browser must be explicitly configured to use it.

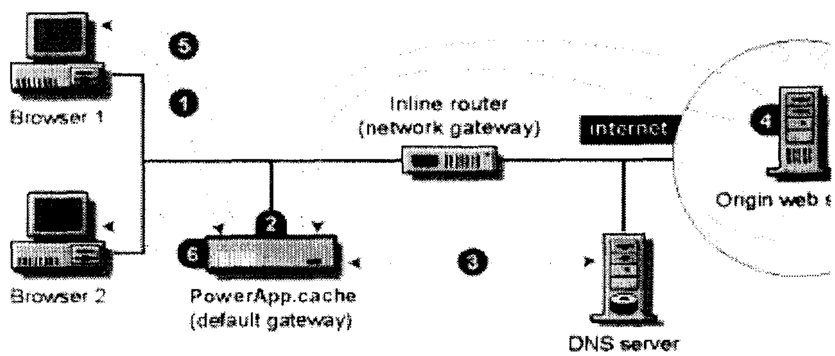


- **Reverse Proxy/Web Server Acceleration** - A reverse proxy is a proxy c sits in front of one or more specific web servers. The cache intercepts all for one or more servers, caches a copy of the objects served, and then s those objects when it next receives requests for them. By serving frequer requested content itself, the cache relieves the origin servers of this load, up processing power on those servers for other tasks, such as managing content requests like e-commerce.

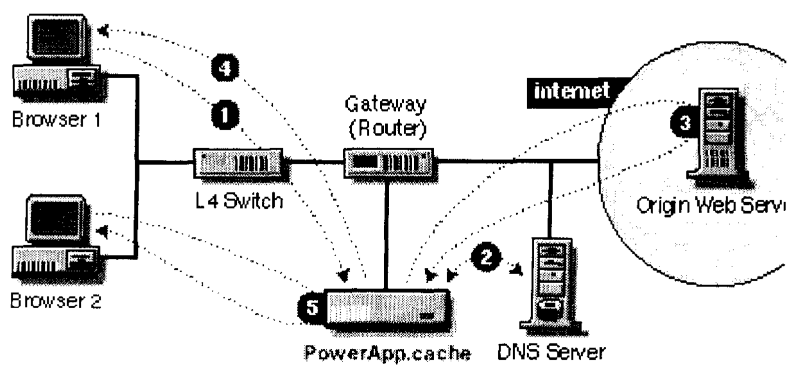


- **Transparent Cache** - A transparent cache is so named because it works intercepting the network traffic transparently to the browser. In this mode, cache short-circuits the retrieval process if the desired file is in the cache. Transparent caches are especially useful to ISPs because they require n set-up modification. The main disadvantage of the transparent approach cache must be placed at a "choke point" in the network through which all to benefit from caching is guaranteed to pass.

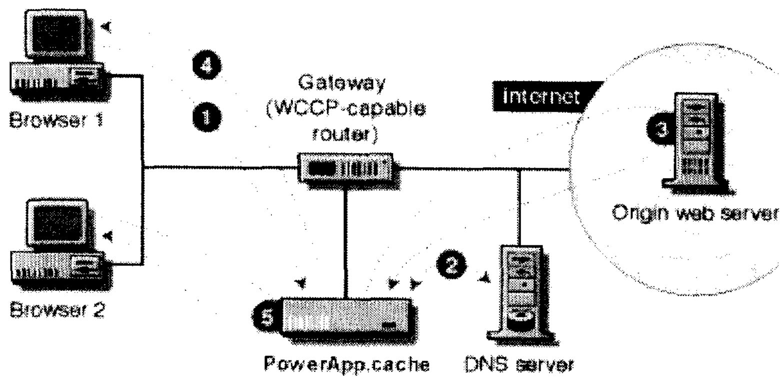
Transparent Proxy - Default Gateway



Transparency with L4 Switch



Transparency with WCCP



Q **8. Who uses web cache solutions today?**

A Both large and small web-caching solutions are in use today. Large web caches are used primarily by service providers and large enterprises or institutions with thousands of users, such as universities and Internet gateway sites around the world. Since so much Internet content is currently located in North America, institutions in Europe, Asia, and Australia use web caching to extract maximum benefit from costly transatlantic links. Small and medium-sized businesses and branch offices benefit from small-scale local web caching. Even sites with less than 20 users can use web caching to improve network response times and reduce WAN link congestion and costs.

Q **9. Where do I deploy cache in a service provider infrastructure?**

A Install a large cache cluster at your main point of access to the Internet. All of your Points of Presence (POPs) will benefit. Client requests are fulfilled at the cache cluster, avoiding traffic on your main Internet access links. To further improve service to clients, you should deploy a cache at each of your POPs. When a client serviced by that POP accesses the Internet, the request is redirected to the POP cache. If the POP cache fulfills the request from local storage, the link to the central access point experiences less traffic.

If the POP cache is unable to fulfill the request, it makes a normal web request. This request, in turn, is routed to the cache cluster at the main access point. If that cache cluster fulfills the request, the traffic on the Internet main access line is avoided, and the client still experiences improved performance. This method is referred to as a cache hierarchy.

Q **10. Where do I deploy cache in my enterprise network?**

A Cache is designed to operate in a hierarchical environment, providing the benefits of caching at every level of your network. Install a cache near the router that is directly connected to the Internet at the top of your network hierarchy to reduce Internet access costs. You can also place a smaller system at remote offices connected to the main office via WAN links to reduce congestion at the central servers and to reduce bandwidth demand on those inter-site links. This hierarchical arrangement is made possible because the system is transparent. Clients do not have to point at any particular cache device to benefit from web caching.

Q **11. What are cache scalability considerations?**

A There are three limitations on scalability of cache: DRAM, Disk storage addressed, and the number of nodes in a cluster. When one cache reaches the limit of its DRAM, storage or processor power, you can usually add more caches to create a cache cluster.

- **DRAM** - The physical RAM in the cache. This ranges from 256MB to several gigabytes.

- **Disk Storage** - Disk access time limits cache performance in two ways: disk operations per second and disk capacity. Even the fastest disk drive is limited to about 100 operations per second (a disk seek followed by a read or write), so when the load on a cache exceeds 100 objects per second, it is necessary to add another disk drive. In terms of capacity, a cache should ideally store every object requested through it for as long as that object is fresh. More traffic means a larger variety of objects requested, which in turn creates a demand for more disk storage.

- **Number of Nodes** - When compared with other applications, caches are very easy to scale. Each cache is a network node, but nodes can be grouped into clusters.

Q **12. What are cache performance considerations?**

A

Cache performance is characterized in three ways: the maximum upstream bandwidth the cache can support, the maximum number of objects per second the cache can handle, or the maximum number of simultaneous TCP/IP connections it can support.

- **Bandwidth** - Caches vary in the speed of upstream network connection they can support. Some caches are designed for operation up to T1 (1.5 megabit) speed, while others can handle T3 (45-megabit) or greater speed, for example. Your bandwidth requirements will depend on the network application you have in mind.

- **Objects/second** - This refers to the cache's processing power. Each cache has a limit on its own ability to process objects regardless of the speed of the connection or the type of objects processed. The range runs from fewer than 100 objects per second on inexpensive caches to thousands of objects per second in carrier-class products.

- **Number of Simultaneous TCP/IP connections** - Caches must handle far more TCP/IP connections than a typical server application. Each browser request represents a TCP/IP connection, and each downstream request to a server is another TCP/IP connection (caches attempt to maintain so-called "persistent" connections to a server they can fetch a number of objects with the same connection). A cache's capacity is therefore limited to the number of simultaneous TCP/IP connections it can support.

Salt Lake City School District

"...In the case of the PowerApp.cache server, it's done everything Dell and Novell promised plus a lot more. We were extremely impressed with the initial product, and the two companies have continued to enhance it, providing great new features such as LDAP [Lightweight Directory Access Protocol] compliance and the ability to log and filter by user name...."
- Sterling Fuhriman, school district computer technician

Background

The Salt Lake City School District
Case Study in .pdf Format (Acrobat File)

The Salt Lake City School District in Salt Lake City, Utah, is responsible for educating nearly 25,000 students in its 37 elementary, middle and high schools. Early on, school officials recognized the value of the Internet to enhance learning and give students essential computer skills that will prepare them for success in the digital economy.

Challenge

In recent years, web access has become increasingly important at the Salt Lake City School District, not only for school business but also for classroom instruction. In response, the district's technical services staff was tasked with providing outgoing Web access for students, faculty and staff, and incoming access for parents and members of the public. This presented the staff with four major challenges.

According to Sterling Fuhriman, a school district computer technician, the first challenge was in meeting performance and availability requirements for classroom instruction. The second challenge was one of prohibiting student access to sites that are not education related-and, in particular, to sites that contain objectionable content. The third challenge was to comply with the Utah governor's mandate to provide parents with online access to student grades through the Web. The fourth challenge was to provide high-quality, fast outgoing and incoming web access without investing a significant amount of money in upgrading the district's network.

Solution

The network's 45 Dell PowerEdge servers currently run Novell® NetWare® 4, and the district plans to migrate to NetWare 5 before the end of the year 2000. The district has also installed five Dell PowerApp.cache servers -- Dell's high-performance, turnkey Internet caching servers based on the Novell Internet Caching System™ (Novell ICS) -- to boost the speed of Internet access.

Today the network includes one server at the district office and one server at each of the four high schools. The high school servers are each connected to the district office server in a hierarchical fashion. Elementary and middle schools are connected directly to the district office server.

The staff has taken advantage of the PowerApp.cache server's filtering capabilities to control student access to objectionable Web sites. The server has built-in filtering, plus it integrates with third-party filtering products such as X-Stop from

Log-on Data. The district has installed X-Stop to take advantage of its extensive filtering list.

Results

The Dell/Novell caching appliance has enabled the school district to provide affordable Internet access to its large user base and still meet its requirements for fast performance, high availability and extensive filtering. *"In terms of hardware,"* Fuhriman notes, *"it costs significantly less to put in the PowerApp.cache to satisfy our need for speed. We've found that we nearly doubled the speed of access to the web server in our district office by accessing it through the cache box-and that's on our local network. With the grade book site, we expect that a single web server front-ended by the cache will handle the expected traffic easily."*

PowerApp.cache is also helping the district keep Internet bandwidth costs in check by minimizing the need to add or upgrade lines to support increased web access. *"We have 37 frame-relay T1 lines connecting all our sites to the District Office"* Fuhriman says. *"The cost to upgrade all of those lines to full T1 or better would have been outrageous."* By storing needed content locally, the cache servers reduce the amount of traffic that must go outside the district office to the UEN cache and to the origin server. That means a lot less traffic over the phone lines, so the existing lines can handle the increased access with ease.

The PowerApp.cache is proving to be a valuable addition to the district's network, providing fast, economical access to the Internet and giving the district the ability to provide high-quality service to parents and the public at an affordable cost. The appliance has far exceeded the district's expectations. *"As a general rule,"* Fuhriman says, *"when we purchase a product, we expect it to do only about 50 percent of what the vendor says it will do. We consider that good. In the case of the PowerApp.cache server, it's done everything Dell and Novell promised plus a lot more. We were extremely impressed with the initial product, and the two companies have continued to enhance it, providing great new features such as LDAP [Lightweight Directory Access Protocol] compliance and the ability to log and filter by user name. Those features will be very beneficial in our environment."*

Attachment C

Before the
Federal Communications Commission
Washington, DC 20554

In the Matter of	)	
	)	
Request for Review of the Decision of the Universal Service Administrator by	)))	
	)	
Cleveland Municipal School District Cleveland, Ohio	))	File No. SLD-190883 Funding Request No. 421840
	)	
Federal-State Joint Board on Universal Service	))	CC Docket No. 96-45
	)	
Changes to the Board of Directors of the National Exchange Carrier Association, Inc.	))	CC Docket No. 97-21

ORDER

Adopted: August 15, 2001 **Released: August 16, 2001**

By the Common Carrier Bureau:

1. The Common Carrier Bureau (Bureau) has under consideration a Request for Review filed by the Cleveland Municipal School District (Cleveland Municipal), Cleveland, Ohio, seeking review of a decision issued by the Schools and Libraries Division (SLD) of the Universal Service Administrative Company (Administrator). The decision under review granted in part Cleveland Municipal's application for Year 3 funding pursuant to the schools and libraries universal service support program.[1] Cleveland Municipal's Request for Review challenges SLD's decision not to grant Funding Request Number (FRN) 421840, which seeks funding for file servers. For the reasons discussed, we deny the Request for Review and affirm SLD's decision.

2. Under the schools and libraries universal service support mechanism, eligible schools, libraries and consortia that include eligible schools and libraries may apply for discounts for eligible telecommunications services, Internet access and internal connections.[2] The Commission's rules require that the applicant make a bona fide request for services by filing with the Administrator an FCC Form 470, which is posted to the Administrator's website for all

[1] Letter from Nathaniel Hawthorne, Esq., on behalf of Cleveland Municipal School District, Cleveland Ohio, to Federal Communications Commission, filed November 13, 2000 (Request for Review).

[2] 47 C.F.R. §§ 54.502, 54.503.

potential competing service providers to review.[3] After the FCC Form 470 is posted, the applicant must wait at least 28 days before entering an agreement for services and submitting an FCC Form 471, which requests support for eligible services.[4] SLD subjects each FCC Form 471 application that it receives to a Program Integrity Assurance (PIA) review and issues funding commitment decisions in accordance with the Commission's rules.[5]

3. On January 20, 2000, Cleveland Municipal filed an FCC Form 471 application seeking Year 3 funding.[6] One of its funding requests, FRN 421840, requested funding for 4 IBM file servers, with a pre-discount cost of $1,700,000.[7] The servers were to provide a number of functions, acting as web servers, hosting the domain name of Cleveland Municipal, and hosting an application known as the "Student Data Warehouse."[8] In its Request for Review, Cleveland Municipal describes this application in following terms: "The student data warehouse servers store data which consists of student records, teacher records concerning student grades, student evaluations, and student addresses, telephone numbers, any discussions with students, parents/guardians about student learning issues, [and] student progress reports."[9]

4. Documentation provided with Cleveland Municipal's FCC Form 471 explains that the need for the database server is to provide a central storage device for this data which teachers from any member school can easily access: "In a large urban district such as Cleveland, there is a very large student mobility. This mobility, (as much as 2000 student changes per day) creates a need for a repository of student information that can be easily accessed by teachers and not dependent on what school the student is enrolled."[10]

---

[3] Schools and Libraries Universal Service, Description of Services Requested and Certification Form, OMB 3060-0806 (FCC Form 470); 47 C.F.R. § 54.504(b); *Federal-State Joint Board on Universal Service*, CC Docket No. 96-45, Report and Order, 12 FCC Rcd 8776, 9078, para. 575 (1997) (*Universal Service Order*), as corrected by *Federal-State Joint Board on Universal Service*, CC Docket No. 96-45, Errata, FCC 97-157 (rel. June 4, 1997), *affirmed in part, Texas Office of Public Utility Counsel v. FCC*, 183 F.3d 393 (5th Cir. 1999) (affirming *Universal Service First Report and Order* in part and reversing and remanding on unrelated grounds), *cert. denied, Celpage, Inc. v. FCC*, 120 S. Ct. 2212 (May 30, 2000), *cert. denied, AT&T Corp. v. Cincinnati Bell Tel. Co.*, 120 S. Ct. 2237 (June 5, 2000), *cert. dismissed, GTE Service Corp. v. FCC*, 121 S. Ct. 423 (Nov. 2, 2000).

[4] 47 C.F.R. § 54.504(b), (c); Schools and Libraries Universal Service, Services Ordered and Certification Form, OMB 3060-0806 (FCC Form 471).

[5] *See* Universal Service Fund Schools and Libraries Program Description for the 2000-2001 Funding Year, issued October, 1999, at 11; *Request for Review by Metropolitan School District of Pike Township, Federal-State Joint Board on Universal Service, Changes to the Board of Directors of the National Exchange Carrier Association, Inc.*, File No. SLD-120821, CC Dockets No. 96-45 and 97-21, Order, 15 FCC Rcd 13891, para. 2 (rel. 2000).

[6] FCC Form 471, Cleveland Municipal School District, App. No. 190883, filed January 20, 2000 (FCC Form 471).

[7] The request includes the actual servers and certain intended components of the servers purchased separately, such as SCSI controllers. *See* FCC Form 471, attachment, "Attachment for E-mail, Web Servers and Student Data Warehouse."

[8] *Id.*

[9] Request for Review at 2.

[10] FCC Form 471, attachment, "Attachment for E-mail, Web Servers and Student Data Warehouse," at 2.

5. On October 13, 2000, SLD issued a funding decision which, *inter alia*, denied FRN 421840 on the grounds that "30% or more of this FRN includes a request for STUDENT DATA WAREHOUSES which is an ineligible product(s)/service(s) based on program rules."[11] Cleveland Municipal appealed this decision directly to the Commission.

6. Applicants may only seek support for eligible services.[12] The instructions for the FCC Form 471 clearly state: "You may not seek for ineligible service, entities, and uses."[13] The instructions further clarify that "[w]hile you may contract with the same service provider for both eligible and ineligible services, your contract or purchase agreement must clearly break out costs for eligible services from those for ineligible services."[14] Although SLD reduces a funding request to exclude the cost of ineligible services in circumstances where the ineligible services represent less than 30 percent of the total funding request, SLD will deny a funding request in its entirety if ineligible services constitute more than 30 percent of the total.[15] An applicant can avoid denial by subtracting out, at the time of its initial application, the cost of ineligible services.

7. File servers are conditionally eligible products. In the *Universal Service Order*, the Commission held that a file server would be classified as a component of internal connections, and thus potentially eligible for discount funding, only if the server "is an essential element in the transmission of information within the school or library."[16] Consistent with this

[11] Letter from Schools and Libraries Division, Universal Service Administrative Co., to Cleveland City School District, dated October 13 2000, at 8 (Funding Commitment Decision Letter).

[12] 47 C.F.R. § 54.504 *et seq.*

[13] Instructions for Completing the Schools and Libraries Universal Service Services Ordered and Certification Form (FCC Form 471) (September 1999) at 18 (Form 471 Instructions).

[14] Form 471 Instructions at 23.

[15] *See Request for Review of the Decision of the Universal Service Administrative Company by Ubly Community Schools, Federal-State Joint Board on Universal Service, Changes to the Board of Directors of the National Exchange Carrier Association, Inc.*, CC Docket Nos. 96-45 and 97-21, Order, DA 00-1517 (Com. Car. Bur. rel. July 10, 2000); *Request for Review of the Decision of the Universal Service Administrator by Anderson School, Federal-State Joint Board on Universal Service, Changes to the Board of Directors of the National Exchange Carrier Association, Inc.*, File No. SLD-133664, CC Docket Nos. 96-45 and 97-21, Order, DA 00-2630, para. 8 (Com. Car. Bur. rel. November 24, 2000). The "30-percent policy" is not a Commission rule, but rather is an SLD operating procedure established pursuant to FCC policy. *See Changes to the Board of Directors of the National Exchange Carrier Association, Inc., Federal-State Joint Board on Universal Service,* CC Docket Nos. 97-21 and 96-45, Third Report and Order in CC Docket No. 97-21 and Fourth Order on Reconsideration in CC Docket No. 97-21 and Eighth Order on Reconsideration in CC Docket No. 96-45, 13 FCC Rcd 25058 (1998). This operating procedure, used during SLD's application review process, enables SLD to efficiently process requests for funding for services that are eligible for discounts but that also include some ineligible components. If 30 percent or less of the request is for funding of ineligible services, SLD normally will issue a funding commitment for the eligible services. If more than 30 percent of the request is for funding of ineligible services, SLD will deny the application in its entirety. The 30 percent policy allows SLD to efficiently process requests for funding that contain only a small amount of ineligible services without expending significant fund resources working with applicants that, for the most part, are requesting funding of ineligible services.

[16] *Universal Service Order*, para. 459.

standard, the Commission found that servers such as "network file servers" were eligible for funding because they were "needed *to switch and route messages* within a school or library."[17] The Commission emphasized that the eligible server's "function is *solely to transmit information* over the distance from the classroom to the Internet service provider . . . ."[18] In a subsequent Public Notice, the Bureau reaffirmed that we support such servers because they are "needed to switch and route messages within a school or library."[19] Thus, to determine whether any file server is eligible for funding as a component of the applicant's internal connections, we look to whether the server is needed as a conduit for information.

8. We find that the Student Data Warehouse servers clearly fail to satisfy that test. These database servers act as the source of content, not as conduits for content which originates elsewhere. Hence, they are not providing "internal connections" as that term was defined in the *Universal Service Order*. Accordingly, we affirm SLD's conclusion that Student Data Warehouse servers are ineligible products.

9. We note that the proposed servers requested by Cleveland Municipal would perform other functions besides acting as host for the Student Data Warehouse. However, we need not determine whether these other functions are fundable. Assuming this is the case, we would still deny the request in full. In the *Universal Service Order*, the Commission held that "schools and libraries may not receive support for contracts that provide only a single price for a package that bundles services eligible for support with those that are not eligible for support. Schools and libraries may contract with the same entity for both supported and unsupported services and still receive support only if any purchasing agreement covering eligible services specifically prices those services separately from ineligible services so that it will be easy to identify the purchase amount that is eligible for a discount."[20] The Commission specifically noted as an example of an improper request the case of an eligible file server which is also "built to provide storage functions to supplement personal computers on the network."[21] That is precisely the case here, and accordingly, under the *Universal Service Order*, the entire request must be denied.

---

[17] *Universal Service Order*, para. 460 (emphasis added).

[18] *Id.* (emphasis added). Another good example of a server necessary to the transport of information are e-mail servers, which act to route e-mail to and from end-users and were upheld as eligible in the pending application. *See*

[19] Public Notice, *Common Carrier Bureau Reiterates Services Eligible For Discounts To Schools and Libraries*, CC Docket No. 96-45, DA 98-1110, 13 FCC Rcd 16570, n.2 (Com. Car. Bur. rel. 1998).

[20] *Universal Service Order*, para. 462.

[21] *Universal Service Order*, para. 461.

10. ACCORDINGLY, IT IS ORDERED, pursuant to authority delegated under sections 0.91, 0.291, and 54.722(a) of the Commission's rules, 47 C.F.R. 0.91, 0.291, and 54.722(a), that the Request for Review filed on November 13, 2000 by Cleveland Municipal School District, seeking review of FRN 421840, is DENIED.

FEDERAL COMMUNICATIONS COMMISSION

Carol E. Mattey
Deputy Chief, Common Carrier Bureau